

SEQUENCE LISTING

<110> Kapeller-Libermann, Rosana

<120> 47174, A NOVEL HUMAN GLYCOSYLTRANSFERASE
AND USES THEREOF

<130> 10448-099001

<140> 09/973,457

<141> 2001-10-09

<150> 60/238,849

<151> 2000-10-06

<160> 6

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2572

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (325)... (2133)

<400> 1

```

gcgcgcgcgc cccgcgcctt cccgcgcgc cccgcgcgc cccgcgcgc ctcaccgctc   60
cccggggcgc ggccgcgcgc tctgagcggg ggatgccggc cgcgcgcgc gacccagcc   120
ccgggcagcc ctctgcgctc tgggggaccc ccggcgccgc tggcccggc cgctgagctg   180
gtgctgaagg gacagctccg gccgagcccc gcagccccc cagcccggg cggtcatgg   240
tccccgaagc cgaagctgaa gccagggccc gggcggggat gctggggatg cccgcgggt   300
gaggcccccg ctgcagcgt gttc atg gcg gtg gcc agg aag atc cga act   351
                Met Ala Val Ala Arg Lys Ile Arg Thr
                        1                      5

```

```

ttg ctg acg gtg aac atc ctg gtg ttc gtg ggc atc gtc ctg ttc tcc   399
Leu Leu Thr Val Asn Ile Leu Val Phe Val Gly Ile Val Leu Phe Ser
  10                      15                      20                      25

```

```

gtg tac tgc cgc ctg cag ggc cgc tcc cag gag ctc gtg cgc atc gtg   447
Val Tyr Cys Arg Leu Gln Gly Arg Ser Gln Glu Leu Val Arg Ile Val
                30                      35                      40

```

```

agc ggc gac cgc cgg gtg cgc agc cga cac gcc aag gtg ggc acg ctg   495
Ser Gly Asp Arg Arg Val Arg Ser Arg His Ala Lys Val Gly Thr Leu
                45                      50                      55

```

```

ggg gac cgt gag gcc atc ctg cag cgc ctg gac cac ctg gag gag gtg   543
Gly Asp Arg Glu Ala Ile Leu Gln Arg Leu Asp His Leu Glu Glu Val
  60                      65                      70

```

```

gtc tac aac cag ctc aac ggc ctt gcc aag ccc atc ggc ctg gtg gag   591
Val Tyr Asn Gln Leu Asn Gly Leu Ala Lys Pro Ile Gly Leu Val Glu

```

75	80	85	
ggg cca gga ggc ctg ggc cag ggt ggc ttg gcg gcc acc ctg cgt gat Gly Pro Gly Gly Leu Gly Gln Gly Gly Leu Ala Ala Thr Leu Arg Asp 90 95 100 105			639
gac ggc cag gag gcg gaa ggc aag tat gag gag tac ggc tac aac gct Asp Gly Gln Glu Ala Glu Gly Lys Tyr Glu Glu Tyr Gly Tyr Asn Ala 110 115 120			687
cag ctc agc gac cgc atc tcc ctc gat cgg agc atc ccc gac tac cgg Gln Leu Ser Asp Arg Ile Ser Leu Asp Arg Ser Ile Pro Asp Tyr Arg 125 130 135			735
ccc aga aag tgc aga cag atg agc tac gcc cag gac ctg ccc cag gtc Pro Arg Lys Cys Arg Gln Met Ser Tyr Ala Gln Asp Leu Pro Gln Val 140 145 150			783
tcc gtg gtc ttc atc ttc gtc aat gag gcg ctg tgc gtc atc ctg cgc Ser Val Val Phe Ile Phe Val Asn Glu Ala Leu Ser Val Ile Leu Arg 155 160 165			831
tcc gtg cac agc gtg gtc aac cac acg ccc tcc cag ctc ctc aag gag Ser Val His Ser Val Val Asn His Thr Pro Ser Gln Leu Leu Lys Glu 170 175 180 185			879
gtc atc ctg gtg gac gac aac agt gac aac gtg gaa ctc aag ttc aat Val Ile Leu Val Asp Asp Asn Ser Asp Asn Val Glu Leu Lys Phe Asn 190 195 200			927
ctg gac cag tac gtc aac aag cgg tac cca gcc ctc gtg aag att gtc Leu Asp Gln Tyr Val Asn Lys Arg Tyr Pro Gly Leu Val Lys Ile Val 205 210 215			975
cgc aac agc cgg cgg gaa gga ctg atc cgc gcg cgg ctg cag gcc tgg Arg Asn Ser Arg Arg Glu Gly Leu Ile Arg Ala Arg Leu Gln Gly Trp 220 225 230			1023
aag gcg gcc acc gcc cca gtc gtc ggc ttc ttt gat gcc cac gtc gag Lys Ala Ala Thr Ala Pro Val Val Gly Phe Phe Asp Ala His Val Glu 235 240 245			1071
ttc aac acg ggc tgg gcc gag ccc gca ctg tgc cgg atc cga gag gac Phe Asn Thr Gly Trp Ala Glu Pro Ala Leu Ser Arg Ile Arg Glu Asp 250 255 260 265			1119
cgg cgt cgc atc gtg ctg cca gcc atc gac aac atc aag tac agc acg Arg Arg Arg Ile Val Leu Pro Ala Ile Asp Asn Ile Lys Tyr Ser Thr 270 275 280			1167
ttt gag gtg cag cag tat gcg aac gcc gcc cat gcc tac aac tgg gcc Phe Glu Val Gln Gln Tyr Ala Asn Ala Ala His Gly Tyr Asn Trp Gly 285 290 295			1215
ctc tgg tgc atg tac atc atc ccc ccg cag gac tgg ctg gac cgc gcc Leu Trp Cys Met Tyr Ile Ile Pro Pro Gln Asp Trp Leu Asp Arg Gly 300 305 310			1263

gac gag tca gca ccc atc agg acc cca gcc atg atc ggc tgc tcc ttc Asp Glu Ser Ala Pro Ile Arg Thr Pro Ala Met Ile Gly Cys Ser Phe 315 320 325	1311
gta gtg gac cgc gag tac ttc gga gac att ggg ctg ctg gac ccc ggc Val Val Asp Arg Glu Tyr Phe Gly Asp Ile Gly Leu Leu Asp Pro Gly 330 335 340 345	1359
atg gag gtg tat ggc ggc gag aac gta gaa ctg ggc atg agg gtg tgg Met Glu Val Tyr Gly Gly Glu Asn Val Glu Leu Gly Met Arg Val Trp 350 355 360	1407
cag tgt ggc ggc agc atg gag gtg ctg ccc tgc tcc cgc gtg gcc cac Gln Cys Gly Gly Ser Met Glu Val Leu Pro Cys Ser Arg Val Ala His 365 370 375	1455
atc gag cgc acc agg aag ccc tac aac aac gac att gac tac tac gcc Ile Glu Arg Thr Arg Lys Pro Tyr Asn Asn Asp Ile Asp Tyr Tyr Ala 380 385 390	1503
aag cgc aac gcc ctg cgc gcc gcc gag gtg tgg atg gat gac ttc aag Lys Arg Asn Ala Leu Arg Ala Ala Glu Val Trp Met Asp Asp Phe Lys 395 400 405	1551
tcc cac gtg tac atg gcc tgg aac atc ccc atg tgc aac cca ggg gtg Ser His Val Tyr Met Ala Trp Asn Ile Pro Met Ser Asn Pro Gly Val 410 415 420 425	1599
gac ttc ggg gac gtg tct gag agg ctg gcc ctg cgt cag agg ctg aag Asp Phe Gly Asp Val Ser Glu Arg Leu Ala Leu Arg Gln Arg Leu Lys 430 435 440	1647
tgt cgc agc ttc aag tgg tac ctg gag aac gtg tac ccg gag atg agg Cys Arg Ser Phe Lys Trp Tyr Leu Glu Asn Val Tyr Pro Glu Met Arg 445 450 455	1695
gtc tac aac aac acc ctc acg tac gga gag gtg aga aac agc aaa gcc Val Tyr Asn Asn Thr Leu Thr Tyr Gly Glu Val Arg Asn Ser Lys Ala 460 465 470	1743
agt gcc tac tgt ctg gac cag gga gcg gag gac ggc gac cgg gcg atc Ser Ala Tyr Cys Leu Asp Gln Gly Ala Glu Asp Gly Asp Arg Ala Ile 475 480 485	1791
ctc tac ccc tgc cac ggg atg tcc tcc cag ctg gtg cgg tac agc gct Leu Tyr Pro Cys His Gly Met Ser Ser Gln Leu Val Arg Tyr Ser Ala 490 495 500 505	1839
gac ggc ctg ctg cag ctg ggg cct ctg ggc tcc aca gcc ttc ttg cct Asp Gly Leu Leu Gln Leu Gly Pro Leu Gly Ser Thr Ala Phe Leu Pro 510 515 520	1887
gac tcc aag tgt ctg gtg gat gac ggc acg ggc cgc atg ccc acc ctg Asp Ser Lys Cys Leu Val Asp Asp Gly Thr Gly Arg Met Pro Thr Leu 525 530 535	1935

aag aag tgt gag gat gtg gcg cgg cca aca cag cgg ctg tgg gac ttc 1983
 Lys Lys Cys Glu Asp Val Ala Arg Pro Thr Gln Arg Leu Trp Asp Phe
 540 545 550

acc cag agt ggc ccc att gtg agc cgg gcc acg ggc cgc tgc ctg gag 2031
 Thr Gln Ser Gly Pro Ile Val Ser Arg Ala Thr Gly Arg Cys Leu Glu
 555 560 565

gtg gag atg tcc aaa gat gcc aac ttt ggg ctc cgg ctg gtg gta cag 2079
 Val Glu Met Ser Lys Asp Ala Asn Phe Gly Leu Arg Leu Val Val Gln
 570 575 580 585

agg tgc tcg ggg cag aag tgg atg atc aga aac tgg atc aaa cac gca 2127
 Arg Cys Ser Gly Gln Lys Trp Met Ile Arg Asn Trp Ile Lys His Ala
 590 595 600

cgg cac tgaccccacc tccgcccgga ccccccacaga cctcgggaag gcgctggggcc 2183
 Arg His

gagccagtgt ggctgagtga cgggggtgtg cccggcagac acagcaggac agggctctat 2243
 gtgcggccag gacagcagag gctgaggggc cgggggtgtg ctgagtgacc aggggtgtcac 2303
 ccaactgcac tggagtacag cttctcctag gacaggcggc tctacccgag ggagggcgctc 2363
 tggggacagt gatgccaaact caaacacgtg ccttctccac ggtatctcct ggccaggctg 2423
 ctgggacagc cgccgcctct gcatgtacca cagccccca cgccccatag ggaggccaag 2483
 ccccggaaca tgcaccaggc tgcaccctgg tgtcttcac ccgcaggcct cccatgctcc 2543
 aagcagcctc ccccagcact tgcggccgc 2572

<210> 2

<211> 603

<212> PRT

<213> Homo sapiens

<400> 2

Met Ala Val Ala Arg Lys Ile Arg Thr Leu Leu Thr Val Asn Ile Leu
 1 5 10 15
 Val Phe Val Gly Ile Val Leu Phe Ser Val Tyr Cys Arg Leu Gln Gly
 20 25 30
 Arg Ser Gln Glu Leu Val Arg Ile Val Ser Gly Asp Arg Arg Val Arg
 35 40 45
 Ser Arg His Ala Lys Val Gly Thr Leu Gly Asp Arg Glu Ala Ile Leu
 50 55 60
 Gln Arg Leu Asp His Leu Glu Glu Val Val Tyr Asn Gln Leu Asn Gly
 65 70 75 80
 Leu Ala Lys Pro Ile Gly Leu Val Glu Gly Pro Gly Gly Leu Gly Gln
 85 90 95
 Gly Gly Leu Ala Ala Thr Leu Arg Asp Asp Gly Gln Glu Ala Glu Gly
 100 105 110
 Lys Tyr Glu Glu Tyr Gly Tyr Asn Ala Gln Leu Ser Asp Arg Ile Ser
 115 120 125
 Leu Asp Arg Ser Ile Pro Asp Tyr Arg Pro Arg Lys Cys Arg Gln Met
 130 135 140
 Ser Tyr Ala Gln Asp Leu Pro Gln Val Ser Val Val Phe Ile Phe Val
 145 150 155 160
 Asn Glu Ala Leu Ser Val Ile Leu Arg Ser Val His Ser Val Val Asn
 165 170 175
 His Thr Pro Ser Gln Leu Leu Lys Glu Val Ile Leu Val Asp Asp Asn

```
<210> 3
<211> 1812
<212> DNA
```

<211> 1812

<212> DNA

<213> Homo sapiens

<400> 3

```

atggcggtgg ccaggaagat ccgaactttg ctgacggtga acatcctggt gttcgtgggc      60
atcgtcctgt tctccgtgta ctgccgcctg cagggccgct cccaggagct cgtgcgcac      120
gtgagcggcg accgccgggt gcgcagccga cagccaagg tgggcacgct gggggaccgt      180
gaggccatcc tgcagcgctt ggaccacctg gaggaggtgg tctacaacca gctcaacggc      240
cttgccaagc ccatcggcct ggtggagggg ccaggaggcc tgggccaggg tggcttggcg      300
gccacctgc gtgatgcagg ccaggaggcg gaaggcaagt atgaggagta cggctacaac      360
gctcagctca gcgaccgcat ctccctcgat cggagcatcc ccgactaccg gccagaaag      420
tgcagacaga tgagctacgc ccaggacctg ccccggtct cctggttctt catcttcgtc      480
aatgaggcgc tgtcggtcat cctgcgctcc gtgcacagcg tggtaacca cagccctcc      540
cagctcctca aggaggtcat cctggtggac gacaacagt acaacgtgga actcaagttc      600
aatctggacc agtacgtcaa caagcggtag ccaggcctcg tgaagattgt ccgcaacagc      660
cggcggaag gactgatccg cgcgcggctg cagggctgga aggcggccac cgccccagtc      720
gtcggcttct ttgatgccc cgctcagttc aacacgggct gggccgagcc cgcactgtcg      780
cggatccgag aggaccggcg tcgcatcggt ctgccagcca tcgacaacat caagtacagc      840
acgtttgagg tgcagcagta tgcgaacgcc gcccatggct acaactgggg cctctggtgc      900
atgtacatca tcccccgca ggactggctg gaccgcggcg acgagtcagc acccatcagg      960
acccagcca tgatcggtg ctccctcgta gtggaccgcg agtacttcgg agacattggg     1020
ctgctggacc ccggcatgga ggtgtatggc ggcgagaacg tagaactggg catgagggtg     1080
tggcagtgtg gcggcagcat ggaggtgctg ccctgtctcc gcgtggccca catcgagcgc     1140
accaggaagc cctacaacaa cgacattgac tactacgcca agcgcaacgc cctgcgcgcc     1200
gccgaggtgt ggatggatga cttcaagtcc cacgtgtaca tggcctggaa catccccatg     1260
tcgaaccagc ggggtggactt cggggacgtg tctgagaggc tggccctgcg tcagaggctg     1320
aagtgtcgca gcttcaagtg gtacctggag aacgtgtacc cggagatgag ggtctacaac     1380
aacacctca cgtacggaga ggtgagaaac agcaaagcca gtgcctactg tctggaccag     1440
ggagcggagg acggcgaccg ggcgatcctc taccctgcc acgggatgtc ctcccagctg     1500
gtgcggtaca gcgtgacgg cctgctgcag ctggggcctc tgggctccac agccttcttg     1560
cctgactcca agtgtctggt ggatgacggc acgggcccga tggccaccct gaagaagtgt     1620
gaggatgtgg cgcggccaac acagcggctg tgggacttca cccagagtgg cccattgtg     1680
agccgggcca cgggcccgtg cctggagggt gagatgtcca aagatgccaa ctttgggctc     1740
cggctggtgg tacagaggtg ctcggggcag aagtggatga tcagaaactg gatcaaacac     1800
gcacggcact ga                                     1812

```

<210> 4

<211> 187

<212> PRT

<213> Artificial Sequence

<220>

<223> Consensus sequence

<400> 4

```

Ser Ile Val Ile Pro Thr Tyr Asn Glu Glu Ala Asp Tyr Leu Glu Glu
 1           5           10          15
Leu Leu Glu Ser Val Leu Ala Gln Ser Thr Leu Glu Asp Ile Glu Ile
 20          25          30
Ile Val Val Asp Asp Gly Ser Glu Thr Asp Glu Thr Val Glu Ile Ala
 35          40          45
Glu Asp Tyr Leu Asp Glu Arg Ile Lys Glu Glu Asn Pro Arg Ile Ile
 50          55          60
Ile Val Ile Arg Leu Glu Glu Asn Ser Gln Gly Pro Ala Ala Ala Arg
 65          70          75          80
Asn Lys Gly Ile Arg Arg Ala Thr Gly Asp Ser Asp Tyr Ile Leu Phe
 85          90          95
Leu Asp Ala Asp Asp Ile Phe Thr Pro Asp Lys Leu Glu Lys Leu Ile

```

```

      100              105              110
Asp Tyr Ala Glu Ala Thr Asp Ala Ala Val Val Leu Gly Ala Ile Asp
      115              120              125
Ala Tyr Glu Tyr Ala Glu Gly Glu Ser Asn Leu Tyr Arg Ile Ala Arg
      130              135              140
Ala Asp Thr Glu Arg Ser Leu Phe Ala Gly Leu Leu Arg Lys Thr Gly
145              150              155              160
Arg Leu Thr Gly Gly Leu Glu Leu Ser Phe Glu Ile Gly Ser Asn Ala
      165              170              175
Ile Tyr Arg Arg Glu Ala Phe Glu Glu Leu Phe
      180              185

```

<210> 5
 <211> 135
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Consensus sequence

```

<400> 5
Arg Gly Tyr Phe Leu Ile Ile Gly Gly Asn Thr Gly Leu Cys Leu Asp
 1              5              10              15
Val Asn Gly Asn Ser Glu Ser Lys Ser Asp Gly Asn Pro Val Gln Leu
      20              25              30
Trp Asp Cys His Gly Gly Gly Asn Gln Leu Trp Lys Leu Thr Tyr Asn
      35              40              45
Glu Ser Asp Gly Ala Ile Arg Ile Asn Ser Asp Leu Cys Leu Thr Val
      50              55              60
Asn Gly Thr Val Thr Leu Tyr Ser Cys Asp Gly Thr Asp Lys Gly Asn
65              70              75              80
Asp Asn Gln Lys Trp Glu Val Asn Lys Asp Gly Thr Ile Arg Asn Pro
      85              90              95
Lys Asn Ser Lys Lys Gly Val Asp Ser Gly Leu Cys Leu Asp Val Lys
      100              105              110
Asp Gly Asn Lys Val Gln Leu Trp Thr Cys Asn Gly Ser Asp Ala Pro
      115              120              125
Asn Gln Lys Trp Ile Phe Glu
      130              135

```

<210> 6
 <211> 149
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Consensus sequence

```

<400> 6
Arg Lys Asp Pro Thr Asp Pro Ile Arg Ser Pro Thr Met Ala Gly Gly
 1              5              10              15
Leu Phe Ala Ile Asn Lys Glu Tyr Phe Glu Glu Leu Gly Thr Tyr Asp
      20              25              30
Pro Gly Met Asp Ile Trp Gly Gly Glu Asn Leu Glu Leu Ser Phe Arg
      35              40              45
Val Trp Gln Cys Gly Gly Arg Leu Glu Ile Val Pro Cys Ser His Val
      50              55              60

```

Gly	His	Val	Phe	Arg	Lys	Arg	Ser	Pro	Tyr	Thr	Phe	Pro	Gly	Lys	Gly	65	70	75	80
Ser	Gly	Lys	Asp	Val	Ile	Ser	Arg	Asn	Thr	Val	Arg	Val	Ala	Glu	Val	85	90	95	
Trp	Met	Asp	Asp	Tyr	Lys	Glu	Tyr	Phe	Tyr	Lys	His	Asn	Pro	Gln	Ala	100	105	110	
Arg	Lys	Val	Arg	Asp	Phe	Gly	Asp	Ile	Ser	Glu	Arg	Lys	Glu	Leu	Arg	115	120	125	
Glu	Lys	Leu	Gln	Cys	Lys	Ser	Phe	Lys	Trp	Tyr	Leu	Glu	Asn	Val	Tyr	130	135	140	
Pro	Asp	Leu	Tyr	Val												145			